GRD Journal for Engineering | Volume 1 | Issue 2 | January 2016
ISSN: 2455-5703

Feature Extraction Techniques for Facial Expression Recognition Systems

Nifi C Joy
PG Student
Department of Computer Science and Information Systems Engineering
Federal Institute of Science and Technology, Mookkannoor P O, Angamaly, Ernakulam, Kerala 683577, India

Dr. Prasad J.C.
Associate Professor
Department of Computer Science and Engineering
Federal Institute of Science and Technology, Mookkannoor P O, Angamaly, Ernakulam, Kerala 683577, India

Abstract

Automatic facial expression recognition has become a progressive research area since it plays a major role in human computer interaction. The facial expression recognition finds its major application in areas like social interaction and social intelligence. However it is not an easy task because the facial image, facial occlusion, face color/shape etc. is not easy to deal with. In this paper, various techniques for feature extraction like Gabor filters, Principal Component Analysis (PCA), Local Binary Patterns (LBP), Linear Discriminant Analysis (LDA), DCT, with different classifiers like Support Vector Machine (SVM) and Neural Networks, which are used to recognize human expression in various conditions on different databases are being examined.
Keywords- Facial expression, Geometric features, Appearance features, PCA, LBP, Gabor, LDA

I. Introduction

A facial expression is a simple physiological activity of one or more parts of the face (eyes, nose, mouth, eyebrows,..) [1]. They were semantically coded with respect to seven basic but "universal" dimensions, i.e., neutral, anger, disgust, fear, joy, sadness, and surprise [2]. Facial expression, being a fundamental mode of communicating human emotions, finds its applications in human-computer interaction (HCI), health care, surveillance, driver safety, animations and computer games, educational software, emotion processing, and deceit detection etc.[3].

A number of novel methodologies for automatic facial expression recognition have been proposed over the last decade. However robust recognition of facial expressions from images and videos is still a challenging task due to the difficulty in accurately extracting the useful emotional features. These features are often represented in different forms, such as static, dynamic, point-based geometric or region-based appearance.

The task of automatically recognizing different facial expressions in human-computer environment is significant and challenging. A variety of systems have been developed to perform facial expression recognition and each system consists of three stages: first, face acquisition; second, facial feature extraction and then facial expression classification [4]. Face acquisition is a preprocessing stage to detect face regions in the input images or sequences. One of the most widely used face detector is the real-time face detection algorithm developed by Viola and Jones. Once a face is detected in the images, the corresponding face regions are usually normalized. Facial feature representation is used to extract a set of appropriate features such as eyes, eyebrow, nose and mouth from original face images for describing faces. It is very essential that only those features should be extracted from images that have highly contribution in expression identification. The final step is facial expression classification that classifies the facial expressions based on extracted relevant features.

Mainly two types of approaches to extract facial features are found: geometry-based methods and appearance-based methods. Geometric feature vectors represent the shapes and locations of facial components by encoding the face geometry from the position, distance, angle, and other geometric relationships between these components. However, geometric feature based methods require accurate and reliable facial component detection, which is difficult to accommodate in many situation . Appearance-based methods apply a single image filter or filter bank on the whole face or some specific region of the face to extract appearance changes. Then, some subspace learning methods, e.g., principle component analysis, linear discriminant analysis are carried out to obtain the subspace representation of the original input. Finally, matching is performed in the learned subspace. But appearance based methods require all the face images have to carefully be aligned and cropped.

II. Facial feature extraction techniques

Facial feature extraction is the process of translating the input data into some set of features. Feature points such as nose, eyes, mouth are extracted and then used as input data to application. Use of feature extraction can help reduce huge amount of data to a relatively small set which is computationally faster. It is influenced by many complications like difference in different pictures of the same facial expression, the light directions of imaging, and the variety of posture, size and angle. Even to the same people, the images taken in different surroundings may be unlike. Various approaches have been proposed to extract these facial points from images or video sequences of faces. The basic approaches are as follows:

A. Geometry Based Techniques

The features measure the displacements of certain parts of the face such as eyebrows or mouth corners. The facial components or facial feature points are extracted to form a feature vector that represents the face geometry. Geometry based method is that expressions affect the relative position and size of various features and that by measuring the movement of certain facial points the underlying facial expression can be determined. The task of geometric feature measurement is usually connected with face region analysis, especially finding and tracking crucial point in the face region.

Ahmad Poursaberi, Hossein Ahmadi Noubari, Marina Gavrilova1 and Svetlana N Yanushkevich[1] introduces Gauss–Laguerre wavelet textural feature fusion with geometrical information for facial expression identification. In this method, both the combined texture and the geometric information of face fiducially points are used to code different expressions. For each input image, the face area is localized first. Then, the features are extracted based on GL filters. Gauss–Laguerre (GL) wavelets are used for texture analysis and the positions of 18 fiducial points represent the deformation of the eyes, eyebrows, and mouth. The combination of these features is used for expression classification. The K-nearest neighbor (KNN) is used for classifying expressions based on closest training examples in the feature space. The advantage of GL filter is its rich frequency extraction capability for texture analysis, as well as being a rotation-invariant and a multi scale approach.

Mingli Song, Dacheng Tao, Zicheng Liu, Xuelong Li and Mengchu Zhou [2] used Image Ratio Features for Facial Expression Recognition Application. Image ratio features combines both the local texture information and the geometric information to recognize facial expressions. Image ratio features effectively capture image intensity changes due to skin deformations. The expression ratio at a point p is defined as the ratio of the intensity of image coordinates of p on the expression face to that of the intensity on the neutral face. Compared with conventional local texture features, e.g., high gradient component, the proposed features are robust to appearance (e.g., albedo) and lighting variations, and can easily be obtained. Given an input face image, first determine whether it is a symmetric or an asymmetric expression by an SVM classifier trained by FAPs. It was demonstrated that image ratio features significantly improve facial expression recognition performance when there are large lighting and albedo variations.

Mohammed Saaidia, Narima Zermi, and Messaoud Ramdani[3] introduced a Multiple Image Characterization Techniques for Enhanced Facial Expression Recognition.. Initially the facial components are extracted. Instead of the matrix of pixels, a reduced representative vector which compact the information needed for processing task is being recommended. Three different types of information is used to enrich the feature vector: Zernike moments, to compact image's geometric characteristics; LBP method which is considered as the most significant way to characterize texture information of the image and DCT transform to obtain its spectral components distribution. The different feature vectors are used separately then combined to train back-propagation neural networks which are used in the facial expression recognition step. A selective tool, called Normalized Mutual Information Feature Selection (NMIFS), is used to compact the combined feature vectors in order to improve the classification performance, optimize the computational cost and reduce the classifier complexity.

B. Appearance Based Techniques

The Features describe the change in face texture when particular action is performed such as wrinkles, bulges, forefront, regions surrounding the mouth and eyes. Image filters are used, applied to either the wholeface or specific regions in a face image to extract a feature vector.

A. J. Calder, A. M. Burton, P. Miller, A. W. Young, and S. Akamatsu [4] introduced the Principal Component Analysis (PCA) method. Each face is treated as a separate one-dimensional array of pixel values. The PCA looks for correlations among the faces (one-dimensional arrays), and where these exist, their coefficients are extracted. The number of components extracted by a PCA referred to as the face's 'signature' and from each signature, it is possible to reconstruct a version of the original face. This is done by weighting each of the Eigen faces with the appropriate component value for the face, and then summing the results. Shape-free method is likely to be of little use to PCA-based models of facial expression recognition. However, PCA faces difficulty if other factors like viewpoint, lighting are varied.

Hong-Bo Deng, Lian-Wen Jin, Li-Xin Zhen, Jian-Cheng Huang [5] uses a New Facial Expression Recognition Method Based on Local Gabor Filter Bank and PCA plus LDA. Gabor filters are band pass filters which are used in image processing for feature extraction. Gabor filters have been found to be particularly appropriate for texture representation and discrimination because their frequency and orientation representations are similar to those of the human visual system. A Gabor filter bank with m frequencies and n orientations is used to extract the Gabor feature for face representation instead of using the entire global filter bank. The method of selecting the LG (mxn) is that the parameter m of frequency increases repeatedly from min to max, and the parameter n of orientation add one for each time. For dimensionality reduction, PCA, an unsupervised learning technique and LDA, a supervised learning method was used. Local Gabor filter bank outperforms global Gabor filter bank in the aspects of shortening the time for feature extraction, reducing the high dimensional feature, decreasing the required computation and storage. When using PCA+LDA method, the dimensionality drastically reduced to 6 dimensions and the recognition performance is improved several percent compared with PCA. However complete elimination of the sensitivity of illumination is a challenge.

Md. Hasanul Kabir, Taskeed Jabid, Oksam Chae[6] introduces A Local Directional Pattern Variance (ldpv) based Face Descriptor for Human Facial Expression Recognition. The LDP feature is a robust facial descriptor and it overcomes the limitations of LBP features since LDP is derived from the edge responses which are less sensitive to illumination changes and noises. Here a new descriptor named the LDP variance (ldpv), which characterizes both the spatial Structure (LDP) and contrast (variance of local texture) information, is introduced. Each LDP code is computed from the relative edge response values in all eight directions, and then, the ldpv descriptor of a facial image is generated from the integral projection of each LDP code weighted by its corresponding variance. The performance of ldpv representation is evaluated with two machine learning methods: Template matching and Support Vector Machine (SVM) with different kernels. Results shows that Ldpv feature is more robust in extracting facial features, and have a superior recognition rate as compared to Gaborwavelet and LBP features. Ldpv also performs stably and robustly over a useful range of low-resolution face images.

Caifeng Shan, Shaogang Gong , Peter W. Mcowan[7] defines Facial expression recognition based on Local Binary Patterns. The operator labels the pixels of an image by thresholding a 3*3 neighborhood of each pixel with the center value and considering the results as a binary number and the 256-bin histogram of the LBP labels computed over a region is used as a texture descriptor. The derived binary numbers is called as the Local Binary Patterns (LBP) and it codify local primitives including different types of curved edges, spots, flat areas, etc. The operator later was extended to use neighborhood of different sizes. Different machine learning techniques, including template matching, Support Vector Machines, Linear Discriminant Analysis and the linear programming technique, are examined to recognize expressions. They also investigated LBP features on low-resolution images, and observe that LBP features perform stably and robustly over a range of low resolutions of face images. However this LBP feature extraction scheme suffers from fixed sub-region size and positions. In order to overcome this limitation, boosting LBP-based classifiers was used, where the distance between corresponding LBP histograms of two face images is used as a discriminative feature, and Adaboost was used to learn most efficient features. A limitation of this work is that the recognition is performed by using static images without exploiting temporal behaviors of facial expressions.

Rania Salah El-Sayed, Prof.Dr. Ahmed El Kholy and Prof. Dr. Mohamed Youssri El-Nahas[8] devised Robust Facial Expression Recognition via Sparse Representation and Multiple Gabor filters. With sparse representation, also known as compressed sensing, each expression can be represented by a set of features, which sufficiently characterize each individual expression. Gabor filters which possess the optimal localization properties in both spatial and frequency domain, is obtained by modulating a sinusoid with a Gaussian. The Gabor wavelet representation of images allows description of spatial frequency structure in the image while preserving information about spatial relations. Using multiple Gabor filters rendered the methd robust to facial expression variations because each filter has specific property to extract. Once the feature extraction is done using the hybrid system comprising of multiple Gabors and Sparse representation, it is trained using SVM classifier. An average recognition rate of (89.28%) was achieved under facial expressions variations.

S L Happy; Aurobinda Routray Routray[9], proposed a novel framework for expression recognition by using appearance features of selected facial patches. A few prominent facial patches, depending on the position of facial landmarks, were extracted which are active during the emotion elicitation. These patches are then processed to obtain the salient patches which contain discriminative features for classification of each pair of expressions. Using the appearance features from the salient patches, the system performs the one-against-one classification task and determines the expression based on majority vote.

III. CONCLUSION

Feature extraction is most important part of face recognition because classification is totally depend on this part. A best feature extraction is not determined without evaluation of face recognition algorithm. That's why best feature set for face recognition are still a problem. This paper discusses various feature extraction technique. Every technique has its pros and cons such as appearance based technique represent optimal feature points which can represent global face structure but disadvantage is high computational cost. However, geometric feature based methods require accurate and reliable facial component detection, which is difficult to accommodate in many situations.

REFERENCES

[1] P. Ahmad, N. A. Hossein, G. Marina, and Y. N. Svetlana, "Gauss– Laguerre wavelet textural feature fusion with geometrical information for facial expression identification," EURASIP J. Image Video Process., vol. 2012, no. 1, pp. 1–13, 2012.
[2] M. Song, D. Tao, Z. Liu, X. Li, and M. Zhou, "Image ratio features for facial expression recognition application," IEEE Trans. Syst., Man, Cybern., Part B: Cybern., vol. 40, no. 3, pp. 779–788, Jun. 2010.
[3] Mohammed S, Narima Z, and Messaoud R:"Multiple Imag Characterization Techniques for Enhanced Facial Expression Recognition",Springer International Publishing Switzerland(2016)
[4] A. J. Calder, A. M. Burton, P. Miller, A. W. Young, and S. Akamatsu, "A principal component analysis of facial expressions," Vis. Res., vol. 41, no. 9, pp. 1179–1208, 2001.
[5] H.-B. Deng, L.-W. Jin, L.-X. Zhen, and J.-C. Huang, "A new facial expression recognition method based on local Gabor filter bank and PCA plus LDA," Int. J. Inform. Technol., vol. 11, no. 11, pp. 86– 96, 2005.
[6] M. H. Kabir, T. Jabid, and O. Chae, "A local directional pattern variance (LDPv) based face descriptor for human facial expression recognition," in Proc. 7th IEEE Int. Conf. Adv. Video Signal Based Surveillance, 2010, pp. 526–532.
[7] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," Image Vis. Comput., vol. 27, no. 6, pp. 803–816, 2009.
[8] EI Sayed, r.s., ei kholy, a., ei nahas "Robust Facial Expression Recognition via Sparse Representation and Multiple Gabor filters", International Journal of Advanced Computer Science and Applications (2013)

[9] S L Happy, Aurobinda R "Automatic Facial Expression Recognition Using Features of Salient FacialPatches", :IEEE TRANSACTIONS ON AFFECTIVE COMPUTING(2015)